Comparative Analysis of Tissue Proteomes for the Discovery of Biomarkers for High Risk of Hepatocellular Carcinoma

D.-J. Yim1, E.-Y. Lee2, J.-A. Lee1, Y.-K. Paik2, H.-J. Wang1, J.-Y. Cheong1, S.-W. Cho1, and K.-B. Hahm1

1AGCG/Ajou University Medical Center, Suwon, Korea; 2BPRC/Yonsei University, Seoul, Korea

In spite of many efforts to find biomarker through the proteomic approach, still many limitations were confronted, including variation between samples, non-disease related physiological or genetic causes, biomarker existing in serum or tissue and comprehensiveness comprising of transcriptional, chromosomal and translational informations. Moreover, since malignancy is arising from multiple steps of inflammation, transformation and dysplasia, we hypothesized that in situ biomarkers can be drawn from comprehensive high throughput analyses of chromosomal genomic hybridization, transcriptional gene expression and proteomic analysis of liver tissues of various clinical stages. For that, we studied 4 healthy HBsAg carriers, 3 chronic hepatitis subjects, 2 liver cirrhosis subjects with portal hypertension and 5 hepatocellular carcinoma subjects, all of which are hepatitis B virus positive. 2DE was performed with IPGphor IEF, image acquisition analysis, tryptic in-gel digestion and MALDI-TOF MS for protein identification were done. 25 spots were found to be significantly associated with the development and progression of chronic hepatitis B disease and 70 spots were identified to be involved in hepatocarcinogenesis. The difference of proteome between non-cancerous and cancerous tissue of the identical person reveals potential candidate proteins related with hepatocellular carcinogenesis. We are doing the comparative analysis with another data set obtained from array CGH and expression analysis, of which results will be shown together to draw plausible biomarkers predicting the progression of liver disease including HCC.

Proteomic Approach to Study the Dynamics of Membrane Proteins in the Olfactory Epithelium of Mus musculus


Ruhr-Universitaet Bochum, Bochum, Germany

The olfactory epithelium (OE) lines the interior surface of the nasal cavity and consists of olfactory sensory neurons (OSNs), basal cells, and supporting cells. When odorants enter the nasal cavity, they dissolve in the mucus covering the luminal surface of the OE and bind to specific olfactory receptors (ORs) located in the plasma membrane of ciliae of OSNs. ORs are G-protein-coupled membrane receptors that are encoded by the largest vertebrate multigene family. Activation of ORs by odorants triggers extensive molecular changes in the cell that lead to short- and long-term neuronal responses. Since the latter response is mainly reflected by changes in protein expression levels, we aim at developing a quantitative MS-based approach to study the OE membrane subproteome from long-term odorant-exposed mice vs. controls.

Low-abundance membrane proteins (MPs) were enriched from murine OE by high salt and high pH treatment followed by in-solution tryptic digestion and nano HPLC/ESI-MS/MS analysis. As a result, we identified 158 unique proteins from which 54% were integral MPs comprising up to 11 transmembrane domains. Evaluation of the subcellular distribution however revealed that plasma membrane proteins (PMPs) represent only a small percentage of all proteins identified. To address this problem and also to allow for the identification of very low-abundance PMPs such as ORs, we adopted subcellular fractionation for selective enrichment of plasma membrane (PM). PM from single murine OE was successfully enriched by sucrose density gradient centrifugation as shown by Western blotting using antibodies against marker proteins of different organelles. In the ongoing work, we apply differential chemical labeling and MS analysis to obtain relative quantitative information on PMPs of the OE from odorant-treated vs. control mice. Displaying changes in the abundance of plasma membrane proteins following long-term odorant exposure is expected to provide us with greater insight into the plasticity of OSNs.
Improved Protein Quantitation Through Peptide Intensity Incorporation

B. Carrillo1, C. Yanofsky1, D. Boismenu2, M. Latterich1, and R. E. Kearney1

1McGill University, Montreal, Quebec, Canada; 2Genome Quebec, Montreal, Quebec, Canada

Isotopic/isobaric labeling strategies are quite common in quantitative proteomics. One key to making quality determinations of differential expression is having accurate measurements of the intensity of the various labeled peptide peaks. There are many processing steps in the analysis of labeled peptides that can distort the resulting calculated ratios and many techniques have been used in attempts to mitigate these errors, from the use of robotics to complete redesign of the labeling reagents. In mass spectrometers with counting detectors, the counting statistics can only be enhanced through the detection of more and more ions. Our data shows that the absolute peak intensity has a significant impact on the quality of the calculated “fold-change” of a peptide, and that this quality must be exploited when combining the individual peptide fold-changes into an estimation of the fold-change of a protein.

In silico proteins with 1 to 15 peptides were generated with random intensities in the empirically determined range of a micro-QTOF mass spectrometer. A second set of peptide intensities was generated for each protein with a randomly generated fold-change in the range of 1 to 35. All peaks were then corrupted by Poisson noise and the fold change estimated via the “standard” peptide-to-protein method (average of the ratios), as well as other algorithms including linear regression, sum of intensities, principal component analysis, and total least squares. Our results show that the latter four methods, which combine peptide intensity information before computing a ratio, perform better in terms of estimating the theoretical ratio than the average of the ratios, the currently used standard. Our study goes on further to show that outlier rejection has a significant impact on the quality of the various fold-change determination schemes.

PSLID and SLIF; Publicly Accessible Proteomics Databases for Subcellular Location

R. F. Murphy

Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Systems biology efforts to build accurate models will require information on the subcellular location of all proteins in any cell type, tissue or organism being modeled. In order to be useful for accurate simulations, this information must consist of more than terms describing subcellular organelles. This is not only because the organelles themselves may vary extensively in size, shape and number between cell types, but also because proteins may be found only within specific portions or subtypes of organelles. Thus in most cases the detailed location of each protein must be determined experimentally. Since fluorescence microscopy is perhaps the most useful method for determining location, our group has developed automated methods for determining subcellular location from microscope images and demonstrated that they can discriminate subcellular patterns better than visual examination. We have created two databases to provide access to information extracted from such images. The Protein Subcellular Location Image Database (PSLID) contains high-resolution images that we and others have collected for the purpose of determining location. It provides tools for clustering proteins to create subcellular location families, and for searching by a number of criteria include similarity to a query image. Perhaps most importantly, PSLID users can download XML files containing generative models for use in cell simulations. These models allow images (distributions) to be generated that reflect the statistical variation in pattern within a particular location family. The Subcellular Location Image Finder (SLIF) performs similar functions for images of proteins automatically extracted from on-line journal articles. It provides structured access to published images depicting subcellular location as an interim solution while comprehensive efforts are underway. Both databases can be queried either interactively or via programmatic interfaces and links are being created between both and the new National Center for Integrative Biomedical Informatics headquartered at the University of Michigan.
Protein Phosphorylation Site Determination Using Microfluidic Chip Interfaced with ETD Ion Trap and QTOF Mass Spectrometry

N. Tang and C. Miller
Agilent Technologies, Santa Clara, California, USA

Posttranslational modifications (PTM) of proteins, such as protein phosphorylation, play important roles in regulation of many cellular functions. The precise determination of the location of phosphorylation is crucial to the understanding of the regulation mechanism. Many biological and analytical methods have been developed to study protein phosphorylation, however, most of them failed to pinpoint the exact location of phosphorylation if multiple phosphorylation sites are present. Electron transfer dissociation (ETD) is a new technique analogous to electron capture dissociation (ECD) where an electron is transferred from a radical anion to a multiply-protonated peptide ion, inducing c and z ions indicative of the peptide sequence. An additional observation for the ETD technique is that the phosphate remains attached to the amino acid during the collision process and allows the specific location of the phosphate group to be observed.

A microfluidic based nanoflow LC interfaced to an ion trap mass spectrometer with ETD, or a QTOF mass spectrometer were used to analyze the samples. Fluoranthene was used as the chemical reagent for ETD. All ETD and CID data were analyzed using Spectrum Mill Protein Identification Workbench.

Preliminary studies were performed on synthetic phosphopeptides to demonstrate the phospho-serine/threonine remained intact in the ETD fragmentation process. Electron transfer dissociation (ETD) is a new technique analogous to electron capture dissociation (ECD) where an electron is transferred from a radical anion to a multiply-protonated peptide ion, inducing c and z ions indicative of the peptide sequence. An additional observation for the ETD technique is that the phosphate remains attached to the amino acid during the collision process and allows the specific location of the phosphate group to be observed.

A microfluidic based nanoflow LC interfaced to an ion trap mass spectrometer with ETD, or a QTOF mass spectrometer were used to analyze the samples. Fluoranthene was used as the chemical reagent for ETD. All ETD and CID data were analyzed using Spectrum Mill Protein Identification Workbench.

Preliminary studies were performed on synthetic phosphopeptides to demonstrate the phospho-serine/threonine remained intact in the ETD fragmentation process. Electron transfer dissociation (ETD) is a new technique analogous to electron capture dissociation (ECD) where an electron is transferred from a radical anion to a multiply-protonated peptide ion, inducing c and z ions indicative of the peptide sequence. An additional observation for the ETD technique is that the phosphate remains attached to the amino acid during the collision process and allows the specific location of the phosphate group to be observed.

A microfluidic based nanoflow LC interfaced to an ion trap mass spectrometer with ETD, or a QTOF mass spectrometer were used to analyze the samples. Fluoranthene was used as the chemical reagent for ETD. All ETD and CID data were analyzed using Spectrum Mill Protein Identification Workbench.

Preliminary studies were performed on synthetic phosphopeptides to demonstrate the phospho-serine/threonine remained intact in the ETD fragmentation process. Electron transfer dissociation (ETD) is a new technique analogous to electron capture dissociation (ECD) where an electron is transferred from a radical anion to a multiply-protonated peptide ion, inducing c and z ions indicative of the peptide sequence. An additional observation for the ETD technique is that the phosphate remains attached to the amino acid during the collision process and allows the specific location of the phosphate group to be observed.

A microfluidic based nanoflow LC interfaced to an ion trap mass spectrometer with ETD, or a QTOF mass spectrometer were used to analyze the samples. Fluoranthene was used as the chemical reagent for ETD. All ETD and CID data were analyzed using Spectrum Mill Protein Identification Workbench.

Preliminary studies were performed on synthetic phosphopeptides to demonstrate the phospho-serine/threonine remained intact in the ETD fragmentation process. Electron transfer dissociation (ETD) is a new technique analogous to electron capture dissociation (ECD) where an electron is transferred from a radical anion to a multiply-protonated peptide ion, inducing c and z ions indicative of the peptide sequence. An additional observation for the ETD technique is that the phosphate remains attached to the amino acid during the collision process and allows the specific location of the phosphate group to be observed.
Efficient Fractionation of Mitochondria and Plasma-Membrane Proteins

K. Steinert, A. Erpenbach, H. Wolgast, P. Porschewski, and C. Heckler
QIAGEN, Hilden, Germany

The complexity of the proteome requires that pre-fractionation of samples is often a prerequisite for efficient analyses. Dividing cells into their structural components often greatly simplifies analysis of pathways and biological processes that are specific to distinct organelles. For meaningful comparisons between samples and other labs, standardized and highly reproducible preparation procedures are required.

Here we present two new kit-based methods for the isolation of mitochondria and fractionation of plasma membrane proteins. Specific enrichment of target fractions was demonstrated by identification of fraction-specific marker proteins after western blotting. Activity tests showed that isolated proteins retained full biological activity.

A test for membrane integrity based on uptake of a carbocyanine dye showed that isolated mitochondria retained their membrane potential. Over 50 proteins in the mitochondria fraction were identified by 2D-PAGE and subsequent MS analysis. Of the proteins identified, over 80% were previously described mitochondrial proteins.

Comparative Analysis of Membrane Glycoproteins from Cultured Cells


Changes in glycosylation have been shown to correlate with changes in the developmental or transformed state of cells. These changes are not restricted to a few proteins, but are global changes seen in many proteins expressed on the cell surface. We describe a new method for analyzing global changes in glycosylation of crude mixtures of glycoproteins. The method enables the direct analysis of crude glycoprotein mixtures without complex sample preparation steps.

The method is based on an array containing more than 20 well-characterized lectins with overlapping specificities. A glycoprotein mixture is biotinylated and applied to the array, and the global glycosylation pattern of the glycoproteins bound to the array is detected with Streptavidin-Cy3. The intensity scan of the array results in a “fingerprint” that is characteristic of the glycan profile of the glycoprotein mixture, and is sensitive to small changes in glycosylation between two samples.

Based on a large dataset of carefully chosen, well-characterized glycoproteins and a large set of membrane protein extracts from different cell lines, a knowledge-based algorithm has been developed to interpret changes in lectin fingerprints.

We present fingerprints and their comparative interpretation from glycoprotein extracts derived from various biological systems (e.g., cells before and after treatment with growth hormones). Results were verified by treating the extracts with various exo- and endoglycosydases, followed by their reanalysis on the lectin arrays.

Protein Expression Analysis in Formalin-fixed, Paraffin-embedded Tissue

K. Steinert, K.-F. Becker, C. Schott, A. Erpenbach, and P. Porschewski

Pathological diagnosis of clinical specimens is mainly performed using formalin-fixed and paraffin-embedded (FFPE) tissues. Formalin fixation is the standard procedure used to preserve morphological details and to prepare tissues in a suitable state for histopathological and immunohistochemical analyses. Because clinical samples represent a comprehensive source of protein expression profiles in different disease states (e.g., before and after therapy), FFPE tissue is valuable for elucidation of differential protein expression patterns associated with disease. Differentially expressed proteins can be used as tumor markers or as targets for new drugs. However, due to the cross-linking of the proteins with formalin, it is difficult to use FFPE tissue for proteomic studies. A prerequisite for the analysis and the comparison of proteins from FFPE tissues is a standardized and reproducible procedure that results in an efficient and quantitative extraction of intact proteins. Here we present a new kit-based method for extracting and quantifying total protein from FFPE tissue sections. The isolated proteins were not degraded and marker proteins such as Her2 could be clearly detected in western blot analysis. Comparison between fresh-frozen and formalin-fixed tissues revealed no differences in protein yield and protein abundances. This technique might be used for the validation and identification of known and novel protein markers in a variety of human diseases.

Bridging the Gap between High-throughput MALDI Screening and Accurate MS/MS Sequence Identifications

M. F. Lopez, S. A. Kuzdzal, D. Sarracino, A. Johnson, and W. Patton

Matrix-Assisted Laser Desorption/Ionization (MALDI) mass spectrometry provides researchers the ability to rapidly screen thousands of samples and quickly discern differentially-expressed biomarker candidates. While MALDI MS remains a driving force in biomarker discovery, making the transition from interesting “profiles” to sequence identifications is typically challenging. Recent advances in sample complexity reduction, MALDI instrument performance and bioinformatics software have enabled researchers to bridge the screening and identification gap. The combination of scaleable, membrane-based fractionation strategies (Vivacience/Sartorius AG), enhanced mass accuracy stability and resolution using MALDI Orthogonal-TOF MS (PerkinElmer SCIEX) and new bioinformatics tools (Nonlinear Dynamics, UK) has resulted in the ability to scale-up interesting biomarker differences discovered from high-throughput MALDI-O-TOF MS screening to facilitate sequence identifications by MS/MS. This process results in accurate sequence information for differentially-expressed, accurate-mass protein/peptide biomarkers. Examples of FT-ICR MS sequence identifications obtained from ovarian cancer serum biomarker MALDI screening studies will be presented.
A Novel Confocal Fluorescence MACROscope for High-throughput Quantitative Imaging of Protein Expression in Cellular Microarrays for Biomarker and Drug-target Discovery

R. S. DaCosta1, E. Lundberg2, P. Constantinou1, A. Asplund3, B. C. Wilson1, F. Ponten3, M. Uhlen2, and H. Andersson2

1Department of Medical Biophysics, Univ. of Toronto, Toronto, Ontario, Canada; 2School of Biotechnology, Royal Inst. of Technology, Stockholm, Sweden; 3Department of Genetics and Pathology, Uppsala Univ., Uppsala, Sweden

Knowledge of the expression levels of proteins in different cells and tissues is an important aspect in proteomics that can contribute to the discovery of novel biomarkers and putative drug-targets. However, to date, this has been limited to low-throughput methods and only qualitative descriptions of protein expression levels. Cell microarray (CMA) technology enables rapid assessment of expression of biomolecular markers by providing the simultaneous analysis of hundreds of tissue and cell specimens. A high throughput imaging technique which provides quantitative information of protein expression at the single-cell-level for large numbers of cancer-derived cell lines or individual patient tissue microarrays would improve our ability to rapidly characterize and compare putative diagnostic and therapeutic "biomarkers." To this end, we have combined the use of a novel confocal hyperspectral fluorescence scanning laser MACROscope™ capable of acquiring images over fields of view up to 2cm x 2cm to simultaneously image approximately 60 different human cancer cell lines (arranged on a single CMA slide) harbouring a variety of molecular abnormalities. The application of cell line microarrays for the assessment of biomarker evaluation was validated by studying the expression of approximately 15 proteins using novel "candidate" mono-specific antibodies. Preliminary results with the MACROscope™ were compared with standard immunohistochemical studies, in vitro fluorescence microscopy and high-resolution multispectral confocal microscopy of immunofluorescent-stained cell pellets. We conclude that the MACROscope™ is a useful instrument for rapid quantitative analysis of expression levels at cellular and subcellular resolutions, and this novel technology may prove useful for biomarker, drug-target discovery and validation processes, as well as permit clinical management of cancer at the individual patient level.
Annotated Spectrum Libraries (ASL) for Sensitive High Speed Protein Identification
D. Fenyo¹, R. Craig², J. Cortens², and R. Beavis³
¹Rockefeller University, New York, New York, USA; ²Beavis Informatics Ltd., Winnipeg, Manitoba, Canada; ³University of British Columbia, Vancouver, British Columbia, Canada

A method for constructing Annotated Spectrum Libraries (ASL) of peptide MS/MS spectra will be described. The methods used to create a set of consensus spectra associated with peptide sequences from approximately 13,000,000 confidently assigned experimental tandem spectra from the Global Proteome Machine Database will be described using a four stage pipeline curation process using the Global Proteome Machine Database was used to improve the collection’s reliability. The reliability of the consensus spectra. The current ASL collection contains data for six model eukaryotic species: human, mouse, dog, cow, rat, and budding yeast. Average ASL gene coverage ranged from 6.9 spectra/gene (human) to 3.8 spectra/gene (cow). The peptide sequences in these libraries have been sequence-aligned with the corresponding ENSEMBL, SWISS-PROT, IPI, or SGD accession numbers, as appropriate. A high speed search engine, X! Hunter, has been constructed to use these libraries, which can identify peptides from sets of experimental mass spectra at a rate of 20,000 spectra/second. The search engine accepts most open tandem mass spectrum formats, e.g., Mascot Generic Format, mzXML, mzData or DTA. The input and output formats are the same as the other two X! series search engines, Tandem and P3. The speed and sensitivity of the search engine will be compared with standard techniques. Application of ASL to the high speed screening of experimental data and instrument control will be discussed. The libraries, protein sequences and all associated software have been made publicly available and they can be downloaded from the project FTP site, ftp://ftp.thegpm.org/projects/xhunter.

Cytochrome P450 Superfamily as a Paradigm for Targeted Proteomics Analysis in Pharmacoproteomics
M. A. Alterman¹ and B. A. Kornilayev²
¹Center for Biologics Evaluation and Research, FDA, Bethesda, Maryland, USA; ²University of Kansas, Lawrence, Kansas, USA

The major objective of personalized medicine is to select individual drug therapies depending upon the correlation of proteomic profiles from diseased tissues with patient response to drug therapy. To a large degree, such response is predetermined by expression profiles of cytochrome P450 isozymes (CYPs). P450s are key enzymes responsible for biotransformation of numerous endogenous and exogenous compounds and are located in almost every tissue. This superfamily is the largest group of enzymes that share high degree of similarity in protein sequence (>3500). Human genome contains 57 CYP genes and 58 pseudogenes. A major gap exists in our knowledge about inter-individual, racial, age, and gender differences in CYP expression on a protein level. Searching and analyzing only or mainly DNA and mRNA information is not sufficient because transcription and particularly translation events not necessarily lead to correlated levels of expressed proteins. None of the traditional methods of CYP isozymes detection that include: selective inhibitors/substrates, antibody-based identification and mRNA analysis can provide reliable quantitative and qualitative information on the individual CYP composition. Moreover these methods require assumption what P450 isozymes are likely to be expressed in a given tissue.

We have developed an integrated proteomic method for the simultaneous qualitative and quantitative analysis of CYP isozyme composition using MALDI-TOF-MS and immunochemistry. Our approach is based on hypothesis that each P450 isozyme possesses unique tryptic peptide(s) (UTP) that could be used for differential analysis of human CYP expression. The analytical method itself provides measurements of the same chemical entity (UTP) in the tryptic digest by two orthogonal analytical techniques—MALDI-TOF MS and immunochemical analysis (ELISA, Western blot, molecular scanner, or extraction with polyclonal anti-peptide antibodies developed against UTP attached to beads). The application of this approach is illustrated on the example of three human CYP isozymes—CYP1A2, CYP2E1, and CYP2C19.
657

Highly Stable Nanoelectrospray Emitters with Integrated Sample Preparation
A. Chonn1, T. Koerner2, R. Xie2, and R. Oleschuk2
1Advanced Integrated MicroSystems Ltd., Vancouver, British Columbia, Canada; 2Queens University, Kingston, Ontario, Canada

We have established a novel mechanism for producing a stable nanoflow (10–1000 nL/min) electrospray ionization. We employ a more robust non-tapered fused silica capillary tip design with an I.D. of 75 μm. A short crosslinked Bead Network column is patterned on the distal end of the capillary. This Bead Network is achieved using porous monolith chemistry to precisely control the crosslinking process at bead to bead contacts and bead to capillary contacts. Importantly, this crosslinking process retains and exposes the native surface of the beads, preserving the separation capacity of the beads. The crosslinked Bead Network column extrudes the fluid at low flow rates and low pressures to generate multiple flow paths, smaller droplets, and multiple “nano” plumes that form a fine “Taylor mist.” We show that this unique architecture has several advantages. The tips are less susceptible to clogging, consistently give higher total ion current (TIC) intensities, and give overall better performance compared to conventional tapered or narrow bore tip designs. Further, the process produces reproducible TICs, with RSD % ≤ 5% for tip to tip variability. Another significant advantage of the crosslinked Bead Network design is the integration of on-line sample preparation with effectively zero dead volume. We show that the short column can effectively function as a solid phase extraction column to neutralize and concentrates samples. These nanoelectrospray emitters will have important applications for high sensitivity, high fidelity mass spectrometry and quantitative proteomics.

658

A Novel “Ion Accounting” Algorithm for Protein Database Searches
G.-Z. Li, D. Golick, M. V. Gorenstein, Johannes P. C. Vissers, J. C. Silva, and S. J. Geromanos
Waters Corporation, Milford, Massachusetts, USA

We have developed a novel, “Ion Accounting” algorithm for protein database searching of parallel LCMS (LCMSE) data. A set of Precursor-Product tables is produced by time-aligning low-energy (precursor) and elevated-energy (product) ions within a retention time tolerance (typically about 1 or 2 second, ~10% FWHM of the chromatographic peak). These Precursor-Products tables are queried against a species-specific, species-specific-reverse, or randomized database with an “Ion Accounting” search algorithm, whereby each time-aligned precursor and product ion is associated with a only one peptide/protein identification. The search is a hierarchal process containing three passes. Pass I matches only correctly cleaved tryptic peptides to the data, returning matches that contain at least one product ion, and whose precursor and product ion mass tolerances are within 5 and 20 ppm, respectively. Each identified peptide is scored and ranked based on its empirical and theoretical physicochemical attributes. The highest-ranked peptides are then collapsed into proteins. The protein having the highest summed-product-ion intensity is considered identified; all precursor and product ions assigned to this protein are removed from further consideration. This cycle is repeated until a certain number of proteins are returned. Pass II then considers the unassigned low and elevated-energy ions and queries them against a subset database of the securely identified proteins found by Pass I. In Pass II, the search criteria is expanded to include the in-source fragments, the neutral losses (H2O and NH3), the missed cleavages and the variable modifications to the Pass I proteins. Pass III then compares the remaining, unidentified ions to the complete database to obtain additional protein identifications (including PMFs). The elimination of the precursor and fragment ions identified by Pass I and II increases the specificity and selectivity of Pass III search.

660

Is Affinity Subtraction of Abundant Serum Proteins Useful and Reproducible?
S. J. Berger, A. B. Chakraborty, G.-Z. Li, C. Dorschel, S. Geromanos, and J. C. Gebler
Waters Corporation, Milford, Massachusetts, USA

A fundamental issue for serum biomarker discovery is the dynamic range of protein concentration. Affinity subtraction of abundant serum proteins is an attractive solution for reducing dynamic range by orders of magnitude, without affecting sample throughput. The analytical benefit of depletion for bottom-up proteomics is the enrichment of lower abundance proteins, as depleted samples are analyzed at the same optimal protein mass load during LC/MS analysis.

In this study, devices for affinity removal of six or twenty abundant serum/plasma proteins were applied to generate depleted human serum samples, in triplicate. Digests of unprocessed and depleted serums were analyzed by a LC/MS methodology where the Q-Tof MS continuously alternated between the acquisition of peptide masses and multiplexed fragmentation data, without precursor ion selection. This was accomplished by directing the beam of ions to the TOF analyzer, while alternating collision cell energies between low and elevated energy states. The resulting data were analyzed by time-resolved mass alignment of detected ions acquired under both energy regimes. Processed data was searched against a human database, and identifications were presumed when a protein was detected by multiple peptides in replicate samples.

A total of 24 data sets (6-Serum, 9-Top6 depleted serum, and 9-Top20 depleted serum were acquired for comparative analysis. It was found that replicate processing of serum by both depletion methods generated consistent protein identifications across sample and experimental replicates, and that targeted proteins were completely or significantly depleted. Interestingly, several intact complement activation pathway proteins targeted for Top20 depletion were quantitatively removed, while biologically active fragments were enriched in depleted serum. Top6 depletion generated ~60% more protein identifications, while Top20 roughly doubled the number of detectable serum proteins. Benchmarking with proteins of known abundance, affinity depletion increased the analytical window for bottom-up proteomics by approximately two orders of magnitude protein concentration.
Producing Qualitative and Quantitative Profiles of Intact Proteins by LC/ESI-TOF MS

I. J. Kass1, K. M. Millea2, J. C. Gebler1, I. S. Krull2, and S. J. Berger1
1Waters Corporation, Milford, Massachusetts, USA; 2Northeastern University, Boston, Massachusetts, USA

Profiling of intact protein mixtures yields information and context not provided by bottom-up (peptide-level) studies. Using intact protein mass information, protein processing steps can be identified, protein modifications (often in combination) can be distinguished, and closely related isoforms can be distinguished. While many labs have studied complex protein mixtures by MALDI-MS, the use of online LC/MS has been limited. This is in large part due to the difficulty of processing the resulting data sets. We have developed a novel approach for the automated processing of intact protein LC/MS data sets, in order to generate qualitative maps of protein mixtures, and to identify quantitative differences between samples. This capability now allows us to address fundamental questions relating to the behavior of complex protein mixtures during LC/MS analysis.

This paper presents systematic analysis of the qualitative and quantitative capabilities of this approach using standard protein mixtures, and cellular lysates. In initial studies we have demonstrated the ability to create qualitative LC/MS protein maps for hundreds of components within a given cell extract, and to retain the chromatographic profile for each of these components. We find that such analyses generates a reproducible data set that is strongly biased toward proteins with masses below 60,000 Da. In more recent studies, the quantitative capacity of these analyses were examined using mixtures of protein standards, and standards spiked into cellular lysates. From these studies, we find that LC/MS analysis of protein mixtures data that are semi-quantitative over approximately two-orders of magnitude intensity. Within this range, both cellular proteins, and spiked protein components demonstrate a "dose-dependent" LC/MS response, from which standard curves can be derived. Ongoing work is aimed at refining data processing conditions to obtain optimal component sensitivity from such analyses, without compromising the ability to quantitate components within the sample.

New Rabbit Hybridoma Fusion Partner (240E-W2), A Novel Tool for a Large Scale of Antibody Development

Y. Ke, R. Pytela, H. Au, Q. Qian, G.-L. Yu, and W. Zhu
Epitomics Inc., Burlingame, CA

Monoclonal antibodies are playing an increasingly important role in understanding of genomic information for protein functions in the post-genomic era. Conventional monoclonal antibody development through mouse hybridoma generation can not keep up with this expanding demand. To develop monoclonal antibodies with a high affinity and wide range of epitopes recognition, Knight and her colleagues (1995) created a rabbit plasmacytoma cell line (240E-1) for rabbit hybridoma generation. However, hybridoma obtained with the original cell line was unstable and no much success was achieved. The founders of Epitomics improved 240E-1 and made it practical to develop monoclonal antibodies in an industrial throughput capacity. But the improved fusion partner (240E-W) contains endogenous rabbit IgG genes derived from the original. Expression and secretion of the endogenous IgG from hybridoma was also shown with antigen affinity purification and protein sequencing of IgG produced from hybridoma. In order to create a new fusion partner without the endogenous IgG, 240E-W was fused with rabbit splenocytes and hybrid cells were screened for IgG negative clones. A new fusion partner with IgG heavy chain negative, named as 240E-W2, was identified by cell ELISA screen and no heavy chain was further confirmed by RT-PCR and genomic PCR. Comparison for rabbit hybridoma generation has shown 240E-W2 performs better than 240E-W in fusion efficiency and positive clone rate. Furthermore, antibodies produced from 240E-W2 hybridoma give better activity as expected. Continuous efforts to create the next generation of rabbit fusion partner without both endogenous heavy and light chains (240E-W3) is on-going in our company.
Improved Detection of Analytes Directly from Whole Blood by Combining Specific Affinity Ligands with Combinatorial Libraries

J. T. Lathrop and Y. J. Hammond
American Red Cross, Rockville, Maryland, USA

Identification of biomarkers and new therapeutics from plasma is a formidable challenge as its complexity prevents evaluation without pretreatment or fractionation. We have developed combinatorially-generated libraries of affinity ligands to capture individual proteins, simultaneously enriching trace and diluting abundant proteins from complex mixtures, and to characterize therapeutic processes and products. This method significantly improves the ability to detect trace proteins; however, certain plasma proteins, irrespective of their abundance, are highly interactive and bind to a disproportionate number of ligands in the library. Unfortunately, depletion of these proteins prior to incubation with the library would remove both these and the trace proteins that bind to them from the sample, decreasing the number of trace proteins that can be detected.

Here we present an improved method for sample preparation that overcomes this problem: a complex mixture like whole blood is co-incubated both with the combinatorial library and a specific affinity resin, identified using other ARC library technology, which binds a highly interactive protein. The highly interactive protein preferentially partitions onto its affinity resin; however, it is not pre-depleted and therefore is not prevented from binding to a portion of the ligands in the library. Furthermore, other proteins that bind to the interactive protein are still present and able to bind to their ligands in the library. Thus, the resolution and enrichment of trace proteins is further improved compared with earlier manifestations of the library sample preparation technology. Furthermore, this system can be used with whole blood, as the polymer support for our libraries is inert to red blood cells and does not activate platelets, coagulation factors, or complement.

In conclusion, we show significant improvement of the library sample preparation technology that can be used with whole blood, improving throughput, decreasing artifacts, preparation time, and costs associated with proteomic sample preparation.

Mining the Serum Fragmentome: Fractionation, Enrichment and Analysis of Carrier Protein Bound Fragments

M. F. Lopez¹, S. A. Kuzdzal¹, D. Sarracino², A. Mikulskis¹, and W. Patton¹
¹PerkinElmer Life and Analytical Sciences, Boston, Massachusetts, USA; ²Harvard-Partners Center for Genetics and Genomics, Cambridge, Massachusetts, USA

Sample complexity reduction is an essential first step for biomarker discovery. Many depletion strategies have been developed to remove the more common, abundant proteins in sera or plasma. Many of these abundant proteins, however, are “carrier” proteins and contain a vast assortment of interesting protein fragments [Mehta et al., Dis. Markers 2003–2004, V19, N1, p. 1–10]. In fact, with an albumin blood concentration exceeding $6 \times 10^{-4}$ M, the probability that even molecules with relatively low binding affinities will be complexed with albumin is greater than 98%. Thus, many potentially interesting biomarkers may be inadvertently lost during protein depletion strategies. This presentation will focus on application of a novel cibacron blue based membrane absorber technology (Vivascience) for the mining of the human serum fragmentome. Multiple fragments representing over 263 unique proteins were fractionated/enriched from human sera using this strategy and identified by FT-ICR MS. Application of a complete biomarker screening and identification platform based on this approach for the discovery of disease-related protein fragments will also be presented.
The HUPO Plasma Proteome Project; Next Phase
G. S. Omenn1, Y.-K. Paik2, and M. Mann3
1University of Michigan, Ann Arbor, Michigan, USA; 2Yonsei Proteome Research Center, Seoul, Korea; 3Max Planck Institute, Munich, Germany

Development and validation of clinically useful plasma or serum protein biomarkers of disease require information about the proteins that can be identified in the circulation in normal individuals. Established in 2002, the HUPO Plasma Proteome Project Pilot Phase evaluated advantages and limitations of many depletion, fractionation, and mass spectrometry technology platforms; compared PPP reference specimens of human serum and EDTA, heparin, or citrate-anti-coagulated plasma; and created a publicly-available knowledge base for others to utilize or collaborate [http://www.bioinformatics.med.umich.edu/hupo/ppp [UM]; www.ebi.ac.uk/pride [EBI]; www.peptideatlas.org/repository [ISB]. After applying an integration algorithm to select a representative protein among a cluster matching exactly to available peptides, the PPP presented and annotated a Core Dataset of 3020 proteins (Proteomics 2005;5:3226). A subsequent paper introduced statistical adjustments for protein length and multiple comparisons testing, generating a stringent subset of 889 proteins (Nat. Biotechnol. 2006, 24: 333). We are the co-chairs for the next phase of the PPP. We will (1) develop standardized operating procedures for specimens, protein and peptide fractionation, and analyses, with attention to replicability of results, to make proteomics practicable for clinical chemistry; (2) select priority PPP proteins for the HUPO Antibody Production Initiative, to generate reagents for biomarker and pathways studies and plasma/organ proteome comparisons; (3) collaborate on informatics, databases, annotations, and error estimation for plasma and serum studies, both HUPO-initiated and published by others; (4) stimulate proteomics technology advances, with special attention to high-resolution/higher-throughput methods and to quantitation of proteins and characterization of modified proteins (primarily glycoproteins and phosphoproteins); and (5) assure paired analyses of plasma and tissue specimens in organ-based and disease-focused proteomics initiatives. [Funded by NIH grant CA 82849, MTTC grant 687, and corporate sponsors].

Proteomic Analysis of Nuclear Protein Complexes That Control Gene Expression
MSKCC, New York, New York, USA

Transcriptional regulation is an elaborate and dynamic process that is dependent on numerous multisubunit protein complexes. Inside the cells, DNA is tightly compacted into highly ordered chromatin structure that impairs the process of transcription. SWI/SNF is one of the nuclear protein complexes with ability to modify chromatin structure in an ATP-dependent manner thus allowing the regulation of transcriptional activation or repression. Our research is focused on the identification of the physical interactions defining SWI/SNF as a functional module of proteins. Systematic identifications of the known subunits of SWI/SNF, together with co-purifying proteins previously unknown as components of the complex, lead to establishment of new function of known and unknown proteins, and ultimately to the establishment of new mechanisms regulating SWI/SNF activity.

For large-scale protein interaction identifications, stable cell lines of HeLa cells expressing FLAG-tagged components of SWI/SNF (Ini1, BRG1, BAF57, and Brm1) were used as a source for affinity purification of the complex. Crude nuclear extracts were loaded directly to immunoaffinity beads constructed with anti-FLAG antibody under native conditions. Binding, wash, and elution were performed under lowest stringency conditions for preservation of possible functional interactions. Captured proteins were analyzed in SDS gel electrophoresis, and subjected to mass spectrometric analysis (gelLC-MS/MS or MALDI) for identifications. Several proteins interacting with SWI/SNF were consistently pulled down with each of the tagged components, among them a protein termed Requiem (Req). All components of SWI/SNF were identified in a reciprocal pull-down using immunoaffinity beads with anti-Req specific antibodies. For further validation of Req as a functional component of SWI/SNF, we performed analysis of the global gene expression in response to treatment with siRNA specific for SWI/SNF components including Req. Co-localization of the complex and Req on promoters of the affected genes in vivo was tested in chromatin immunoprecipitation (ChIP) analysis.
A New Method of Generating Decoy Peptides in SEQUEST to Quantify False-Positive Rates for Peptide Results

H. Ko, K. Soo, J. Candlin, and D. Chiang
Sage-N Research, Inc., San Jose, California, USA

A new method has been developed to generate decoy peptides for Sorcerer-SEQUEST searches of tandem MS data, with the goal of quantifying the false-positive rate for peptide results. Existing methods of generating decoy peptides usually involve concatenating a target and a reversed decoy protein database and then performing a virtual digestion of the resulting database, followed by a search against the resulting target and decoy peptides. The new scheme—theorized by S.P. Gygi—generates decoy peptides at search-time by retaining the amino acid residues at the N- and C-termini for each target peptide, while reversing the residues in between.

There are a number of advantages to this approach. The ratio of target to decoy peptides searched is always exactly one-to-one, because every target peptide has a decoy peptide pair with an identical mass; this is not true for a concatenated database search. The computational overhead for the reverse peptide search is also substantially lower—only 1.3× slower than the equivalent non-decoy search, compared to 2× slower for a concatenated database search. Additionally, this approach allows researchers to preserve their original protein databases, thereby separating algorithmic enhancements from data.

Decoy peptides identified during the search are clearly labeled in the output, allowing any downstream analysis tools to handle them appropriately. Validation was performed via comparisons to the current de-facto method of concatenated reversed databases, and show that the resulting false-positive rates are comparable.

A Workflow for False-Positive Analysis of Peptide Results Using a Novel Implementation of SEQUEST for the Trans-Proteomic Pipeline

K. Soo, H. Ko, and D. Chiang
Sage-N Research, Inc., San Jose, California, USA

A method has been implemented to allow the existing Trans-Proteomic Pipeline (TPP) tools from the Institute of Systems Biology (ISB) to handle decoy peptides for high-confidence protein identifications without requiring the presence of a concatenated database (consisting of a target protein database and its reversed decoy pair). For each decoy peptide found as a top result in a spectrum, two things are done. The first is a modification of its associated SEQUEST XCorr value to be substantially lower, allowing the TPP tools to more ably apply its stochastic models to distinguish between incorrect and correct peptide identifications. The second is to associate the peptide with a randomly chosen protein from the target protein database. The assignment of the decoy peptides to these random proteins distributes the incorrect identifications in such a way that no particular protein becomes artificially promoted from an incorrect to a correct identification. We demonstrate that these changes have the desired effect in TPP’s post-search validation tools, and that the false-positive rate can thus be determined from our TPP-dependent tools.
Proteomic Analysis of Chaperones and Stress Proteins in *Halobacterium* sp. NRC-1

H. D. Shukla

University of Maryland Biotechnology Institute, Baltimore, Maryland, USA

*Halobacterium* sp. NRC-1 is an extremely halophilic archaeon and has adapted to optimal growth under conditions of extremely high salinity. Its proteome is highly acidic with a median pI of 4.9, a unique characteristic which helps the organism to adapt high saline environment. In the natural growth environment, *Halobacterium NRC-1* encounters a number of stressful conditions including high temperature and intense solar radiation, oxidative and cold stress. A modified 2-D gel electrophoretic procedure, employing IPG strips in the range of pH 3–6, enabled improved separation of acidic proteins. Combining experimental data from 2-D gel electrophoresis with available genomic information, allowed the identification of at least 30 cellular proteins involved in many cellular functions: stress response and protein folding (CctB, PpiA, DpsA, and MraA), DNA replication and repair (DNA polymerase A α subunit, Orc4/CDC6, and UvrC), transcriptional regulation (Trh5 and ElfA), translation (ribosomal proteins Rps27ae and Rps6 of the 30 S ribosomal subunit; Rpl31ae and Rpl18e of the 50 S ribosomal subunit), transport (YufN), chemotaxis (CheC2), and housekeeping (ThiC, ThiD, FumC, ImD2, GapB, TpiA, and PurE). To study the heat shock response of *Halobacterium NRC-1*, growth conditions for heat shock were determined and the proteomic profiles under normal (42°C), and heat shock (49°C) conditions, were compared. Using a differential proteomic approach in combination with available genomic information, bioinformatic analysis revealed five putative heat shock proteins that were upregulated in cells subjected to heat stress at 49°C, namely DnaJ, GrpE, sHsp-1, Hsp-5 and sHsp-2. Constitutive expression of stress proteins and chaperones help the organism to adapt and survive under extreme salinity and other stress conditions. The upregulated expression pattern of putative chaperones DnaJ, GrpE, sHsp-1, Hsp-5 and sHsp-2 under elevated temperature clearly suggests that *Halobacterium NRC-1* has a sophisticated defense mechanism to survive in extreme environments.

Analytical Strategy for Biomarker Discovery in Plasma using Fractionation by Chromatography followed by MicroRotofor and 2D Gels

S. Freeby, A. Paulus, A. Posch, T. Wehr, N. Liu, and K. Academia

Bio-Rad Laboratories, Hercules, CA

Biomarker discovery programs using whole blood, plasma or serum are pursued in clinical research laboratories, the pharmaceutical industry and proteomics core labs among others. Plasma and serum are relatively easy to obtain and store and therefore large depositories exist for specific diseases. The problem however is high abundance of relatively few proteins exhibiting more than 90% of the protein mass of a serum or plasma sample. Recently, immuno-depletion strategies, which remove the 6, 12 or 20 most abundant proteins, have been investigated. Typically, these methods work with only 100 μl of serum or plasma, limiting both any additional separation steps and ultimately the detection limit and accessibility of low abundant proteins. While serum and plasma covers a concentration range of 10 orders of magnitude, immuno-depletion can access only the low pg/ml or high ng/ml range. While digging deeper into the proteome allows the detection and identification of more protein species, it also increases the changes of finding differential expression between a disease and control sample. Therefore methodologies using multiple separation steps are being explored although they lack the elegance and simplicity of one-step procedures. Here, we are describing a methodology using 4 individual protein separation steps. Starting with 3.5 ml plasma, we remove in a chromatographic step 65% of the protein load by eliminating albumin and IgG’s. The remaining 35% plasma proteins are fractionated twice with the MicroRotofor, first with a broad pH range of 5 to 8 and subsequently with a pH range of 1 pH unit into 10 fractions. The resulting MicroRotofor fractions are further analyzed with Experion, the automated protein-sizing-on-a-chip method, or with 2D gels. We are comparing the virtual gels of a MicroRotofor and Experion with the results of a 2D gel in a diabetic sample.
Plasma Protease Substrate Profiling of tPA Treatment in Acute Ischemic Stroke Patients: A Step Toward Pharmacoproteomics at the Bedside

M.-M. Ning1, D. A. Sarracino2, F. Buonanno1, B. Krastins2, S. Chou1, D. McMullin3, and E. H. Lo1

1Massachusetts General Hospital, Harvard Med. School, Boston, Massachusetts, USA; 2Harvard-Partners Center for Genetics and Genomics, Boston, Massachusetts, USA; 3New York University, New York, New York, USA

Tissue plasminogen activator (tPA) is the only FDA approved medical therapy for acute ischemic stroke. As a serine peptidase, tPA has been shown to affect the expression of other proteases and cell signaling in the brain. In animal models, our group and others have shown that tPA increases expression of matrix metalloproteinases (MMPs), which are implicated in the breakdown of the blood-brain barrier (BBB) leading to intracranial hemorrhage (ICH), a dangerous side-effect of tPA. However, studies in human subjects are scarce.

We perform proteomic substrate profiling, or degradomics, in human plasma to provide a global picture of the effect of tPA on cell-cell signaling in the brain in the context of acute stroke pathophysiology.

Methods: Plasma from acute stroke patients was analyzed pre and post intravenous (IV) tPA using zymography, ELISA, and LC/MS/MS (LTQ-FT) to identify proteases and substrates of interest. Non-tPA-treated patients and healthy controls, matched by age and risk factors, were also analyzed within the same time frame.

Results: Pre- and post-tPA plasma samples from the same patient demonstrate distinct degradomic patterns, with substrates of specific proteases found among the degraded proteins. Within the same time frame, non-tPA-treated acute stroke patients did not show the degradomic changes characteristic of tPA treatment. Control patients without strokes had little change in their degradomic profiles over time.

Conclusion: Our findings suggest that tPA treatment affects the plasma degradomic profiles of acute stroke patients. By studying the substrate profile of tPA treatment in acute stroke, we hope to demonstrate the feasibility of performing pharmacoproteomics at the bedside to elucidate the mechanisms of thrombolysis-related therapeutic efficacy. Further studies are under way to validate these findings.

Identification of Chlamydia Peptides Presented by Major Histocompatibility Complex by Nano-LC/MS/MS

University of British Columbia, Vancouver, British Columbia, Canada

The aim of this study was to identify MHC-binding peptides derived from the intracellular microbe Chlamydia muridarum using an immunoproteomic approach and murine dendritic cells (DC) infected with C. muridarum. Peptides were isolated by immunoprecipitating loaded MHC molecules and then separated from larger proteins by ultrafiltration through a 5 kDa molecular weight cut-off membrane. Peptides were then concentrated using STAGE tips prior to LC/MS/MS analysis with an LTQ-FT and an LTQ-Orbitrap. Fragment spectra were searched against a combined mouse and C. muridarum database using Mascot with no enzyme specificity. Because of the lack of specific terminal sequences in MHC-bound peptides the high mass accuracy provided by the two Fourier transform mass spectrometers used here was crucial for confidently assigning peptide identifications. Optimization of sample preparation and LC/MS/MS conditions allowed for identification of several peptides confirming the processing and presentation of Chlamydia proteins by the MHC class I and II pathways. Sequence information on the MHC-associated peptides unique to infected cells will provide valuable information for development of vaccines since the class II MHC-peptide complex plays a critical role in cell-mediated adaptive immunity against Chlamydia.
Integrated Proteomic and Metabonomic Analyses of the Cardioprotective Effects of Nitrite Administration in the Heart

Boston University School of Medicine, Boston, Massachusetts, USA

Nitrite has recently been shown to offer protection from ischemia-reperfusion injury in the heart, liver and brain, and to have signaling molecule-like properties, affecting the enzymatic activities of soluble guanylyl cyclase and cytochrome P450 and the expression levels of heat shock protein 70 and heme oxygenase-1. Intrigued by these findings, we have used an integrated proteomics and metabonomics approach to further characterize the impact of changes in systemic nitrite availability on the heart.

Cardiac tissue was harvested from male Wistar rats 5 min. to 24h after systemic administration of physiologic and therapeutic levels of sodium nitrite. In addition to the assessment of cellular redox status, the concentrations of NO-related metabolites were immediately determined by gas phase chemiluminescence and HPLC. Crude mitochondria and post-mitochondrial cytoplasmic supernatants were prepared by differential centrifugation and subjected to 2D-PAGE analysis. Differentially expressed proteins were identified by in-gel trypsin digestion, MALDI-TOF MS, and peptide mass fingerprint analyses.

Our metabonomic and redox analyses revealed that systemic administration of nitrite induced short-term spikes in the levels of cardiac S-nitroso, N-nitroso, and heme-nitrosyl species, followed by a large and persistent increase in the ratio of oxidized to reduced ascorbate. These changes were accompanied by significant alterations to the cardiac proteome, involving proteins with known functions in energy metabolism, redox balance, chaperone activity, cell structure, contractility, and nitric oxide metabolism, including both up- and down-regulation of protein expression, and changes in post-translational modifications.

This integrated proteomic and metabonomic approach is a step toward elucidating the scope and mechanism of cardioprotection and the potential physiological activity of nitrite.

Acknowledgements: This project was funded by NIH grants P41-RR10888 (to C.E.C.), S10-RR15942 (to C.E.C.), HL69029 (to M.F.), and NHLBI contract N01-HV-28178 (to C.E.C.).

Comparative Proteomics; Post-Translational Modifications in Sickle Cell Disease

M. E. McComb, C. Dauly, A. Ohiambo, D. H. Perlman, H. Huang, M. H. Steinberg, H. W. Farber, E. S. Klings, and C. E. Costello
Boston University School of Medicine, Boston, Massachusetts, USA

The vascular pathology of sickle cell disease (SCD) is characterized by altered nitric oxide (NO) metabolism and increased oxidant stress. Pulmonary hypertension (PH) of SCD is an important risk factor for mortality; its etiology is unclear and likely multi-factorial. We hypothesize that, in PH of SCD, oxidative stress results in post-translational protein modifications that contribute to the pathogenesis of SCD.

We utilized a proteomic approach to study post-translational protein modifications in the plasma of patients with PH of SCD. Plasma, obtained from age, sex and racially-matched Sickle Cell Disease (SCD) patients with pulmonary hypertension (PH) and healthy controls, were albumin-depleted, then were separated by two-dimensional HPLC on a Beckman PF2D™. Differential 2D expression maps were generated with the Beckman software ProteoVue™ and DeltaVue™, and fractions showing appreciable differences in protein expression levels, pI, or reversed-phase retention time, were targeted for MS analyses. Albumin was recovered and analyzed by MS.

We identified several abundant and medium-abundance proteins that appear to vary with SCD and PH, and we have identified several protein post-translational modifications which also correlated to the disease phenotype. Modifications of transferrin, an important molecule in iron metabolism and transport, and apolipoprotein A1, a molecule required for HDL-mediated activation of endothelial nitric oxide synthase, were observed and could play a role in the pathogenesis of PH in SCD, or could serve as important biomarkers of disease. Differential plasma proteomics such as this study may have a significant impact on the detection and treatment of SCD, PH and other human diseases.

This project was funded by NIH grants P41-RR10888 (to C.E.C.), S10-RR15942 (to C.E.C.), U54 HL070819 (to M.H.S.), and NHLBI contract N01-HV-28178 (to C.E.C.).
Membrane Protein Characterization in 3-Dimensions via 2-D HPLC and Mass Spectrometry

C. Dauly, D. H. Perlman, C. E. Costello, and M. E. McComb
Boston Univ. Sch. of Medicine, Boston, Massachusetts, USA

Rapid identification and characterization of proteins from complex samples is a challenging goal, especially for membrane proteins whose hydrophobicity makes 2D-SDS-PAGE analyses difficult. We evaluated the capabilities of a 2-dimensional HPLC protein fractionation system based on size exclusion chromatography (SEC) followed by reversed phase (RP) chromatography coupled with mass spectrometry (MS) for the characterization of membrane proteins obtained from endothelial cells. Endothelial cells were harvested in a sucrose buffer and subjected to ultracentrifugation. The pellet constituting the membrane fraction was washed at high pH to remove membrane associated proteins and de-lipidated by methanol/chloroform extraction. The resulting proteins were dissolved in a detergent buffer to constitute the membrane fraction and separated by SEC and RP chromatography. Our 2-D HPLC methodology yielded orthogonal separation and gave high peak capacity for the membrane proteins. The UV traces from the RP separation were used to construct protein maps using ProteoVue™ software (Beckman Coulter). Post separation, the proteins in aliquots of the liquid fractions were directly characterized as intact entities by MS, enabling further resolution of several proteins: hence the 3rd dimension of separation was provided by MS. The molecular masses of the intact proteins varied from 8 kDa (mitochondrial ATP Synthase-[epsilon] chain), to 71 kDa (HSP70). Differences between theoretical and experimental masses were attributed to the presence of isoforms, truncation and post-processing/post-translational modifications. The proteins were then digested by trypsin and peptide mass fingerprint (PMF) analysis was performed using MALDI-TOF MS, allowing correlation of the intact protein masses with the PMF results. Overall we demonstrated that the 2D-HPLC MALDI-TOF MS system allows complex protein samples to be rapidly fractionated and identified by integration of the information on intact molecular weights and peptide mass database searches.

This project was funded by NIH grants P41-RR10888, S10-RR15942, and NHLBI contract N01-HV-28178 (to C.E.C.).

TPR-Mediated Protein-Protein Interaction of the Cytochrome C Maturation System in Escherichia coli

D. Han, K. Kim, J. Oh, and Y. Kim
College of Medicine, Seoul National University, Seoul, Korea

Post-translational maturation of the C-type cytochrome involves covalent attachment of heme to the apocytochrome polypeptide by two thioether bonds. In Escherichia coli, the attachment of the cofactor requires eight or nine specific proteins, which are organized in a membrane protein complex.

The nrf genes (NrfABCDEFGH) are required specifically for the maturation of a special C-type cytochrome, NrfA, involved in the periplasmic formate-dependent nitrite reduction. NrfG protein, composed of three tetratricopeptide repeat (TPR) motifs, was predicted to interact with other nrf proteins via TPR domain. TPR domain, which is composed of three tetratricopeptide repeat (TPR) motifs, was predicted to interact with other nrf proteins via TPR domain. TPR domain, which is composed of 3–16 tandem-repeats of 34 amino acid residues, binds specific peptide ligands and are thought to mediate protein-protein interactions in a variety of biological systems.

This report represents that NrfG, belonging to heme lyase (NrfEFG), is required to fulfill a specialized role in the assembly of a functional formate-reductase complex, and interacts with c-type cytochrome (NrfA). The interaction mode of NrfG with NrfA is a newly identified TPR-mediated interaction in the complex formation, compared with the peptide-binding modes of the Hsp70/Hsp90 organizing protein and the PEX5 receptor which have been reported to date. To elucidate the interaction mode, structural approaches along with biochemical assay were carried out to confirm the interactions of NfrG with other nrf proteins. Finally, we would propose that NrfG interacts with other Nrf proteins, especially C-type cytochrome (NrfA), by TPR-mediated specific interaction where a new binding pattern of the loop-helix motif was utilized in the interaction between TPR domain and partner molecule.
Here we present two software applications developed for proteomics analysis of mass spectrometry (MS) data. BUDSS (Boston University Database Search Shell) is a GUI based program designed to aid the processing and management of various MS data files. BUPID (Boston University Protein Identification) is a probability-based peptide mass fingerprint (PMF) database search engine. Linked together they form the basis of an automated MS data analysis platform being developed within our laboratory at BUSM.

BUDSS performs two separate but related functions: automatic conversion of processed MS data files (includes both mzXML and mzData formats) into several common formats accepted by different software applications and the submission of these converted data files to several web based database searching engines (Mascot, Ms-Fit, BUPID). For data management, the software supports multiple users/projects and organizes data and settings individually for each user/project. All these functions are accessible through a user-friendly graphical interface.

BUPID applies a log-likelihood ratio calculation to determine the probability that a protein is present within the sample based on the interrogation of PMF data. The model distinguishes the null hypothesis: a set of peaks in the spectrum is generated by the random background, from the alternative hypothesis: the same set of peaks is generated by peptides corresponding to a specific protein. Peak assignment is based on the log-likelihood ratio as opposed to matching peaks with peptides within the mass tolerance. The mass spectrum is first matched with unmodified protein sequences. The top scored proteins in the search result are then used to construct a database that includes PTMs and amino acid variants, against which the spectrum is searched again. Results are ranked by their probability scores.

This project was funded by NIH NHLBI contract N01-HV-28178 (to C.E.C.).

Age-related macular degeneration (AMD) is the most common cause of legal blindness in people over 60 years of age, and is estimated to affect millions of individuals worldwide. At present there is no fundamental cure for AMD, although some success in attenuating choroidal neovascularization has been obtained with surgical excision or photodynamic therapy. Major clinical characteristics of AMD, where atrophy occurs around the choriocapillaris with clear boundaries. The accumulation of debris-like material referred as drusen between the retinal pigment epithelium and Bruch’s membrane is observed to precede this exudation and atrophy. Early detection is most needed for the early treatment of AMD. There has been a strong demand to find AMD-specific proteins in plasma. Here we present a comparative proteome analysis of human plasma obtained from both 6 AMD and 6 age-related cataract (AC) patients. Proteins were extracted from plasma after the depletion of abundant human serum albumin (HAS) and IgG. Proteins were digested to peptides by trypsin. Tryptic peptides were subjected to nano-electrospray ionization (nano-ESI) MS, and MS/MS in profile mode following reverse-phase (RV) micro-liquid chromatography (micro-LC) separation. Semi-quantitative comparison was performed by the DeCyder MSTM Software, and protein identification was carried out by the MASCOT™ search engine. In this study 9 proteins were upregulated specific to AC, and those specific to AMD were only 1 protein called contactin3 although these proteins are needed to be validated. Details of analysis and its future direction would be discussed.
A Comparison of Immunodepletion Strategies for Plasma Biomarker Discovery

Australian Proteome Analysis Facility (APAF) Ltd., Sydney, New South Wales, Australia

Plasma has long been considered the most reliable source for protein and peptide biomarkers as it provides “snapshots” of the human phenotype at particular intervals—reflecting a true representation of differences between health and/or disease. The extraordinary dynamic range of proteins in plasma (i.e., between highly abundant and disease specific biomarkers) can be as high as $10^{12}$, severely limiting biomarker discovery potential on proteomic platforms that struggle to distinguish protein concentration greater than 4–5 orders of magnitude (e.g., large format 2DE-DIGE analyses using pi 4–7 IPG strips maximally loaded with Agilent MARS column depleted plasma).

Since detection of biomarkers is limited by the amount of protein that can be loaded onto proteomic arrays (e.g., 2DE/DIGE, MD-LC/MS, iTRAQ/MS), strategies for the removal of abundant plasma proteins have been developed by several companies. This study compares many of these approaches and the data demonstrates their utility in achieving some increase in sensitivity and detection of some less abundant plasma proteins.

For the first time, we show the use of a new parallel processing polyclonal IgY approach for the removal of high abundance proteins reliably and cheaply. This method is called CAPI (cyclic abundant protein immunodepletion). CAPI demonstrates that prior fractionation of plasma using novel methodologies that separate plasma into groups of abundant proteins prior to immunization into chickens can produce IgY that when chicken plasma (i.e., between highly abundant and disease specific biomarkers) can be loaded onto proteomic arrays (e.g., 2DE/DIGE, MD-LC/MS, iTRAQ/MS), strategies for the removal of abundant plasma proteins have been developed by several companies. This study compares many of these approaches and the data demonstrates their utility in achieving some increase in sensitivity and detection of some less abundant plasma proteins.

Protein nanoarrays can serve as a useful platform for ultraminiaturized bioanalysis. In many cases, protein samples are extremely limited and temporally unique. Therefore, one can get the many advantages to create the protein chip for high-throughput analysis using protein nanoarray technology. In this work, protein nanoarrays with integrin, BSA and antibody were fabricated on prolinder coated Au surface by dip-pen nano lithography (DPN). Protein was arrayed at 80% humidity condition with 0.01 nN contact force. We investigate protein-protein interaction on nanoarray using two different methods. The first method is measuring the height change of nanoarrayed spots after interaction between integrin and vitronectin that is a specific ligand of integrin. And the second method is measuring the interaction force between nanoarrayed capture protein and target protein tightly immobilized on cantilever surface. These results suggest that protein nanoarrays can be used to analyze protein-protein interaction at single molecular level.

Methods for Protein-Protein Interaction on Protein Nanoarrays

$^{1}$Hoseo University, Asan, Rep. of Korea; $^{2}$Chungbuk National University, Cheongju, Rep. of Korea

Protein nanoarrays can serve as a useful platform for ultraminiaturized bioanalysis. In many cases, protein samples are extremely limited and temporally unique. Therefore, one can get the many advantages to create the protein chip for high-throughput analysis using protein nanoarray technology. In this work, protein nanoarrays with integrin, BSA and antibody were fabricated on prolinder coated Au surface by dip-pen nanolithography (DPN). Protein was arrayed at 80% humidity condition with 0.01 nN contact force. We investigate protein-protein interaction on nanoarray using two different methods. The first method is measuring the height change of nanoarrayed spots after interaction between integrin and vitronectin that is a specific ligand of integrin. And the second method is measuring the interaction force between nanoarrayed capture protein and target protein tightly immobilized on cantilever surface. These results suggest that protein nanoarrays can be used to analyze protein-protein interaction at single molecular level.

Determination of the P-STM Antibody-Recognizable Phosphorylation Sites on Lamins A and C in Mitotic HeLa Cells by Mass Spectrometry

J.-S. Yu, K.-Y. Chien, and C.-w. Ho
Chang Gung University, Tao-Yuan, Taiwan

P-STM, a phosho-epitope-specific antibody generated against the auto phosphorylation/regulatory site of p21-activated kinase (PAK2) (antigen sequence: SKRSpTMVGTPY), not only recognizes the autophosphorylated/activated PAK2 but also many phosphoproteins including lamins A/C in mitotic HeLa and AA31 cells (Yu et al., Biochem. J. 334, 121–131, 1998; Tsai et al., J. Cell. Biochem. 94, 967–981, 2005). Lamins A/C are known to be phosphorylated by p34cdc2 kinase in mitotic cells (Heald et al., 1990). The P-STM recognizable epitopes on lamins A/C may serve as good targets to monitor lamins A/C phosphorylation levels during cell cycle. Lamins A/C immunoprecipitated from unsynchronized HeLa cells could be phosphorylated in vitro by p34cdc2 and this phosphorylation created epitope(s) on lamins A/C to P-STM, suggesting that the phosphorylation site(s) mediated by p34cdc2-catalyzed reaction in vitro may represent the one(s) recognized by P-STM in mitotic cells. The aims of this study are to identify the P-STM recognizable epitopes on lamins A/C in mitotic phase and to quantify the sites of lamins A/C differentially phosphorylated at interphase and mitotic phase. We applied mass spectrometry and molecular biology technologies to identify and quantify lamins A/C phosphorylation site(s) recognized by P-STM. We show that Thr-19 of lamins A/C phosphorylation after phosphorylation by p34cdc2 kinase represents the sole epitope to P-STM. The different levels of lamin A/C peptide 125GAQASS19TLP25R in control and nocodazole-treated cells revealed by SILAC (stable isotope labeling amino acid cell culture) method suggest that Thr-19 of lamins A/C is differentially phosphorylated at interphase and mitotic phase.